R E P O R T R E S U M E S

ED 003 896 48

SOME NOTES ON BILINGUAL LEXICOGRAPHY.

BY- HARRELL, RICHARD S.
INDIANA UNIV., BLOOMINGTON

REPORT NUMBER NDEA-VI-16-11 PUB DATE 60
EDRS PRICE MF-$0.09 HC-$0.56 14P.

DESCRIPTORS- *DICTIONARIES, *LEXICOGRAPHY, *BILINGUALISM, *ARABIC, *CONFERENCES, BLOOMINGTON, INDIANA

BASIC PROBLEMS IN THE COMPOSITION OF BILINGUAL DICTIONARIES WERE DISCUSSED, INCLUDING (1) DETERMINING USER-GROUPS FOR WHOM A DICTIONARY IS INTENDED, (2) INDICATING MULTIPLE SEMANTIC MEANINGS, AND (3) COMPILING A DICTIONARY FOR USER-GROUPS WHO AIM AT EXPRESSING THEMSELVES IN A LANGUAGE AS WELL AS COMPREHENDING IT. FURTHER DISCUSSIONS DEALT WITH ARABIC LEXICOGRAPHICAL PROBLEMS AND A METHODOLOGY FOR DEVELOPING ARABIC DICTIONARIES. THIS REPORT IS ONE OF A SERIES OF 13 PAPERS PRESENTED AT A CONFERENCE ON LEXICOGRAPHY, INDIANA UNIVERSITY, NOVEMBER 11-12, 1960. (RS)

ED003896

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

16-11

SOME NOTES ON BILINGUAL LEXICOGRAPHY

Richard S. Harrell
Georgetown University

A meeting such as ours is not the place for a detailed technical discussion either of general lexicographical methodology or of the specific details involved in the construction of a bilingual dictionary for any given pair of languages. Witness is the fact that time taken up with external consultants and with staff meetings to iron out the details of our Arabic lexicographical problems at Georgetown University has already exceeded the time allotted to this conference. Our colleagues working in the fields of Greek and Pashto lexicography surely have a similar story to tell.

It seemed more profitable, and more likely to be of interest to this group, to fasten upon a few specific problems of general lexicographical interest and then to show, grosso modo, how these problems were met in the Arabic lexicographical work now in progress at Georgetown. The points offered for consideration are: For speakers of what language is the bilingual dictionary intended? Is the bilingual dictionary to concern itself primarily with written or with spoken material? Is the bilingual dictionary primarily to serve as an aid to comprehension or to expression? Finally, an appendix is added outlining the organizational routine of the lexicographical work in progress at Georgetown.

I. Three General Problems

A primary problem in the composition of a bilingual dictionary is to decide whether the work is intended principally for the speakers of X-language or the speakers of Y-language. It is clearly impossible to pay equal attention to both X-speakers and Y-speakers in one and the same work.

Points of conflict already arise at the preliminary point of deciding what items are to be entered in the dictionary. Suppose that X is English and that our dictionary is intended primarily for English speakers. Given

this, there is no good reason why such an English word as circumnavigate should be included. The English speaker already knows what circumnavigate means, and to learn how to express its meaning in Y-language he can look up such periphrases at to sail around or to go around. If, however, our dictionary is intended primarily for Y-speakers, there is no guarantee that the Y-speaker will not somewhere or other encounter circumnavigate, and he would need to be able to look it up.

Fortunately, this is one point on which some double service is possible: Theoretically, it does not hurt the English speaker to have circumnavigate included in the alphabetical English word list. Practically, however, there actually is some disadvantage to the English speaker in having circumnavigate included in the dictionary. In the real world no dictionary is ever written without some a priori limitation on size and cost. With a limitation on size, the need of Y-speakers to have such an entry as circumnavigate in an English-Y dictionary can only be given limited attention, and then only after the contrary needs of the English speaker have been given fullest consideration.

On the other hand if the compilers of an English-Y dictionary have primarily in mind the needs of Y-speakers, entirely other considerations will prevail. Since no countries other than the United States and Canada indulge in football, an English-Y bilingual dictionary oriented solely to the needs of English speakers would be wasting space to include the word tackle in its sporting sense.* But, for the Y-speaker who is to have any broader cultural contact with the American scene, tackle would be an appropriate dictionary entry. The same is true of numerous cultural words which a Y-speaker would encounter on the scene of the English speaking world but for which an English speaker would not expect to find equivalents

---

* Canadian French is a possible exception to this specific item.

in Y-language. One has only to think of such items as forty-niner, mugwump, and chile con carne to perceive the extent of the problem. The item chile con carne (usually just chile) can serve to illustrate the pitfalls involved. There is no Spanish equivalent for this item. The food preparation referred to was invented in Texas and is unknown in Mexico. The etymologically equivalent Spanish words chile con carne have only the literal meaning "red pepper with meat", and the average monolingual waiter in a Mexican restaurant would probably be greatly puzzled to hear a customer request such a thing.

The above considerations are all stated in the context of an alphabetical English word list with Y-equivalents. Exactly the same considerations apply, mutatis mutandis, to an alphabetical list in Y-language with English equivalents.

Beyond the level of vocabulary selection there are two further main points where a decision must be made as to whether the dictionary is designed for primary utility to X-speakers or Y-speakers. In what language are general directions to be given, not merely in the introduction but in individual entries, and in what way are multiple meaning equivalents to be handled? In neither case is a really fruitful compromise possible, at least not in the real world.

A bilingual introduction to a dictionary is easy enough to arrange, but to have such indices as "irregular plural", "imperative", "noun", etc. given bilingually in every entry is beyond the pale of practical manageability. There is no middle ground on this issue. Either X-speaker or Y-speaker must be discriminated against at the expense of the other. The most nearly successful attempt that I have seen at favoring speakers of both languages in a bilingual dictionary is in the German-English, English-German Dictionary of Everyday Usage*, In cases where an entry has multiple

---

* J. Alan Pfeffer, ed. New York: Henry Holt & Co., 1951.

glosses, an illustrative sentence, in both English and German, is given for each gloss. Such an ideally helpful approach, however, is likely to be possible only in the case of dictionaries of relatively modest scope.* A dictionary aiming at a broader lexical coverage would be made too expensive and too bulky by such an approach.

The problem of indicating multiple meanings is a manifestation at the semantic level of the familiar phenomenon that different languages are self-contained systems exhibiting only limited isomorphism with one another. What does the lexicographer do when confronted with one X-word for which there are two or three or perhaps five or six Y-equivalents, each one with its own unique shade of meaning and not interchangeable with the others? An example may be taken from a readily available contemporary Spanish-English, English-Spanish dictionary.** The Spanish word pera is glossed

> pear; goatee; sinecure, easy job...

and the English word hot is glossed

> caliente; caluroso; cálido; picante (como el pimentón chile, ají, etc.); furioso; fresco, reciente;...

These two items were taken at random and are excellent illustrations of the point at issue. It is obvious that the gloss of pera is entirely in favor of the English speaker. A Spanish speaker seized by an urge to know the English equivalent of pera would receive only limited help from this gloss. He will certainly know that pear, goatee, and sinecure are each the equivalent of some shade of meaning of pera, but he has no way of telling which of the three English glosses refers to a fruit, to chin whiskers, or to bogus employment. Presumably, with an exercise of ingenuity and consider-

---

* Each section of the Pfeffer work contains fewer than 10,000 head-words.

** Castillo, Carlos and Otto F. Bond, Spanish-English, English-Spanish Dictionary (New York: Pocket Books, Inc.)

able cross-referencing, the Spanish speaker could solve this problem, but the dictionary does not immediately do it for him.

The English word list, on the other hand, is slanted in favor of the Spanish speaker. A Mexican looking up hot would receive rather comprehensive information about its Spanish equivalents, but the American who simply wished to know how to say hot in some given context is confronted by a bewildering string of equivalents. He presumes they are not mutually interchangeable, but he is given no clue as to what the semantic range of each is.

The Castillo and Bond dictionary thus arranges its English-Spanish section for maximum utility to Spanish speakers while arranging the Spanish-English section for maximum utility to English speakers. It would have been equally possible to favor either the Spanish speaker or the English speaker in both parts. A bilingual dictionary which goes both ways, as the Castillo and Bond dictionary does, is in a sense two works in one, and the solution they have adopted is a common one. That is, to arrange the glosses of the alphabetical word list for the benefit of the foreign rather than the native speaker.

The implied assumption of the approach illustrated by the Castillo and Bond dictionary is that the English speaker will mostly look up Spanish words and that the Spanish speaker will mostly look up English words, and scarcely at all vice versa. The generalization of the assumption is that the user of a bilingual dictionary is primarily concerned with understanding a foreign language and scarcely at all with expressing himself in a foreign language.

And what is this assumption but a disguised reassertion of our traditional academic approach to foreign languages, an approach which emphasizes reading for comprehension and which scarcely conceives of any other possible goals in foreign language education? The needs of contemporary foreign

language education are more involved, both in the United States and in the rest of the world. More and more people have an active need to express themselves in foreign languages, both in writing and in speaking. What of the American in Japan, who, although expecting a monolingual Japanese friend, has to go out unexpectedly? He might like to do such a human thing as tack a note on his door. And what of the American working in an Egyptian village who needs to talk about things for which he does not know the Arabic equivalents? In the past the American would either have had no monolingual Japanese friends or he would have had an interpreter-translator to help him along. Likewise there would have been little chance, in the past, of finding an American in an Egyptian village. In the world of today more and more Americans are abroad, and they are having to do both their own talking and their own writing. Fortunately a fair number of them show signs of actually wanting to do it in addition to merely having to do it. The traditional bilingual dictionary is poorly adapted to their needs. The "For-speakers-of-what-language?" problem of the bilingual dictionary is doubled in cases where there is a serious gap between spoken and written forms of a language. The question faced is "For speech or for writing?" The problem is doubled again if the dictionary maker is faced with the practical limitations of being able to do either an X-Y or a Y-X dictionary but not both. The question then faced is, "Which is to be stressed, comprehension or expression?" By now it is clear that to speak of composing a bilingual dictionary of X language and Y language is a misapprehension. There are various possible varieties of bilingual dictionaries of X and Y languages.

II. Arabic Lexicographical Problems

The Georgetown Arabic Research Program had to contend with all three issues discussed above. The decision was to compose primarily with the needs of the English speaker in mind, to favor speech over writing, and to

favor expression over comprehension. The Georgetown schedule accordingly calls for the preparation of an English-Arabic dictionary in each of three modern Arabic dialects, Moroccan, Syrian, and Iraqi. The allotted work period is two years, from June, 1960 to June, 1962, and the goal is medium sized dictionaries of maximum everyday utility for the educated layman. At present there are no English-Arabic dictionaries at all in these dialects, not even poor ones.

Many of the problems peculiar to Arabic lexicography have their roots in the tangled soci-linguistic phenomena of the well known dichotomy between Classical Arabic and the various colloquial dialects. Briefly put, the socio-linguistic situation of the Arab world is today roughly parallel to that of the Latinized provinces of the Roman empire just before the emergence of the modern romance languages as publicly accepted media of written communication and serious intellectual culture.

There are several immediate consequences of this diglossia* for the maker of a colloquial Arabic dictionary. To begin with he must work almost solely with oral material. Such usual sources of lexicographical raw materials as magazines, novels, etc. are not available for him. And how is his material to be transcribed? There is no conventialized spelling to be followed. The lexicographer must invent his own.

Inventing a spelling for an Arabic dialect is not a trivially simple task. What is to be done about a host of foreign loan words which are often only imperfectly assimilated to Arabic phonology? And in the case of many loan words which are perfectly assimilated at an unsophisticated level, educated Arabs use the pronunciation of the source language and consider the assimilated pronunciation vulgar. Furthermore Arabic morphophonology is intricate, and a strictly phonemic transcription of any

---

* CF. Charles A. Ferguson, "Diglossia", Word, 15.325-340, 1959.

Arabic dialect is never satisfactory. A separate decision as to whether and how much a morphemic as opposed to a strictly phonemic transcription must be made on each of a large number of separate points, and no two people ever quite agree on every one of them.

But even before he can permit himself the luxury of any of the above considerations, the lexicographer of colloquial Arabic has to decide exactly whose speech he is writing a dictionary of. There is a gross reality to such terms as "Moroccan Arabic" and "Iraqi Arabic", just as there is in such terms as "British English" and "American English", but there are many internal differences within each of these macro-varieties. In Egypt, despite great local diversity, the dialect of Cairo is a fairly normalized educated standard, spoken and understood everywhere within the country. The three "macro-dialects" which are the subject of the Georgetown dictionary project, Moroccan, Syrian, and Iraqi, present far more painful problems. For Iraqi the educated speech of Baghdad has been taken as the base for work, and for Moroccan the educated speech of Casablanca-Rabat. But in neither Morocco nor Iraq is there an exclusively accepted colloquial standard throughout the entire country.

"Syrian Arabic" is an even more complex matter. The spoken Arabic of historical geographical greater Syria (i.e. the area now divided among the political units of Syria, Lebanon, Israel, and Jordan) is a palpable linguistic unity, but, village dialects aside, there is no metropolitan educated standard for the entire area. Of the three metropolitan centers of Damascus, Beirut, and Jerusalem, there is no compelling sociological or linguistic reason for choosing one over the others. The Georgetown solution was to make the English-Syrian dictionary basically an English-Damascene dictionary and to include Beirut and Jerusalem forms wherever these differ from Damascene usage. It would have been equally possible to use either the speech of Jerusalem or Beirut as a base. Our decision to use Damascene as our point of departure was determined by the simple non-linguistic fact that our Syrian staff is most directly familiar with it.

APPENDIX

On August 29-30, 1960, Miss Ann M. Driscoll of the G. and C. Merriam Company and Professor W. Freeman Twaddell of Brown University visited the Georgetown University Arabic Research Program as consultants. The following outline on methodology was prepared for Georgetown by Miss Driscoll, and its suggestions have been adopted. The references to "Pfeffer" in the outline refer to the German-English, English-German Dictionary of Everyday Usage*, the alphabetical English word list of which was adopted as a point of departure for the English-Arabic dictionaries being compiled at Georgetown.

---

* J. Alan Pfeffer, ed., op. cit.

Methodology for Arabic-Dictionaries Project

Ann M. Driscoll
G. and C. Merriam Company

Submitted: August 30, 1960

I. Keep everything relating to all three dictionaries on colored 3 x 5 slips, giving the following significance to each color:

1. green - definitions from Pfeffer and others
2. pink - manuscript for Iraqi
3. blue - manuscript for Moroccan
4. yellow- manuscript for Syrian
5. white - comment slips, query slips, supplementary slips, AA slips

II. To prevent any confusion of colors, give the supply of slips for his color to each editor and let no one else have access to them.

III. Cut and paste Pfeffer on the green slips, starting ahead of the place where editors are now working.

1. Paste an entire entry in sequence, continuing on a second slip, if necessary.
2. Paste each phrase on a <u>separate</u> slip, with the entry word written at the top of the slip, keeping the phrases in alphabetical order under the entry word, thus:

   get

   get around.....

3. Paste as close to the top left-hand corner of the slip as possible and leave about 1/2 inch at the bottom of each slip.
4. Write <u>cont</u> at the bottom of every continued slip and "entry word - cont" at the top of the next slip. Use 1, 2, etc. after "cont" at the top of the next slip if there is more than one continued slip.

IV. Set these green slips up in a filing cabinet with removable drawers as the beginning of a file into which every scrap of material used for or in all three dictionaries will be filed. In each drawer leave room for expansion.

V. In the file keep the slips in this order:

1. definitions from Pfeffer and other dictionaries (green)
2. Iraqi manuscript (pink)
3. Moroccan manuscript (blue)
4. Syrian manuscript (yellow)
5. comment slips, etc. (white)

VI. As soon as enough material has been gathered in three drawers to warrant it, assign a drawer to each of the three editors. Each will then edit the drawer by writing his entries on slips of the color assigned to him and inserting them in the file. The entry should be handled as follows:

1. a single sense to each slip, with all of the "heading" material (as entry word, inflected forms, feminine form) either on the same slip with sense 1 or on a separate slip depending on its length.
2. cont at the bottom of each contined slip.
3. the entry and cont (with 1, 2, etc. following when necessary) and 1, 2, etc. (sense numbers) preceding when necessary.
4. These slips should be prepared with the idea that they will be used as copy for the printer, which means that all typography and punctuation should be clearly indicated. Typography is indicated by underscoring as follows:

   roman - no underscoring

   boldface - ~~~~~

   italic - _____

   small caps - =====

5. Perhaps no systematic proofreading need be done until the galley stage.

6. Every slip handled by an editor should be stamped with his name and the date. This may be done by the editor himself as he goes along or by a clerk boy as the editors finish with them.

7. The Pfeffer and similar green slips are to be used as guides only, not a copy.

8. At the end of the editing, slips of each color will be extracted from the file and taped to 8½ by 11 sheets for the printer.

VII. As an editor thinks of a comment he wishes to make about a word anywhere ahead of the edited part of the alphabet, he makes out a white slip and puts the word in the upper left corner, the comment, his name and the date in the lower right corner and drops the slip in an "out" box on his desk. Such slips will be picked up periodically and filed at their own alphabetical places in the main file.

VIII. When an editor wishes to put a comment slip in the section of the alphabet on which he is working, he merely inserts it unless it is a question to someone else. If it is a question to someone else that he needs answered immediately, he writes in the upper right of the slip the name of the person to whom it is addressed and drops the slip in the "out" box. It is then delivered, answered directly on the slip or on an attached slip, dropped in the "out" box of the answerer, and thus returned to the questioner. The questioner then writes "closed" across the top of the slip and sees that it gets filed in its proper alphabetical place as "backing." If it is something he wishes to call to the attention of an editor who will be working on the drawer after, he leaves the slip in its alphabetical place in the drawer.

IX. When an editor comments on a term that has already been edited, the comment is to be <u>filed in a supplementary file</u>, which will not be reviewed until A-Z has been edited. It is the responsibility of the file clerks to see that <u>nothing new</u> is added to a section that has already been edited.

X. When an editor needs a policy decision, he writes the general heading (as Plurals) in the upper left and AA in the upper right. These slips should then be delivered to Mr. Harrell who either writes the answer immediately or has the slip filed for the staff meeting. Whether answered immediately or in a meeting, every policy decision whould be recorded on 3x5 white slips in copies enough for each editor to have one at his desk.